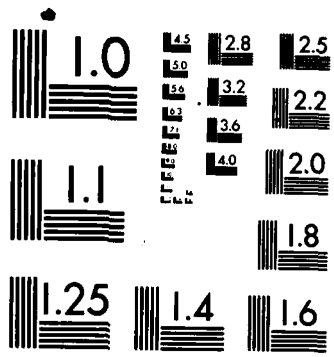
AD-A161 978 NEW CONFIDENCE INTERVAL ESTIMATORS USING STANDARDIZED TIME SERIES(U) CORNELL UNIV ITHACA NY SCHOOL OF OPERATIONS RESEARCH AND INDU D GOLDSMAN ET AL 1/1

UNCLASSIFIED DEC 84 TR-J-84-15 N00014-81-K-0037 F/G 12/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AD-A161 978

SCHOOL OF INDUSTRIAL AND
SYSTEMS ENGINEERING
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332

TECHNICAL REPORT NO. J-84-15
December, 1984

NEW CONFIDENCE INTERVAL ESTIMATORS USING STANDARDIZED TIME SERIES

by

David Goldsman

and

Lee Schruben

This research was partially supported by the Office of Naval Research under contract N00014-81-k-0037 at Cornell University. Lee Schruben is on the faculty of the School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853.

Abstract

We develop new confidence interval estimators for the underlying mean of a stationary simulation process. These estimators can be viewed as generalizations of Schruben's so-called *standardized time series area* confidence interval estimators. Various properties of the new estimators are given.

In this note, we present new confidence interval estimators for the underlying mean μ of a stationary simulation process. These estimators can be viewed as generalizations of the so-called *area* estimators given in Schruben (1983) and Goldsman (1984).

Consider the stochastic process $X_1,\ldots,X_m$. Define

$$\bar{X}_j \equiv \frac{1}{j}\sum_{k=1}^{j} X_k \ , \quad j=1,\ldots,m,$$

$$\sigma^2 \equiv \lim_{m\to\infty} m\mathrm{Var}(\bar{X}_m) \ , \quad \text{and}$$

$$T_m(t) \equiv \frac{\lfloor mt\rfloor(\bar{X}_m-\bar{X}_{\lfloor mt\rfloor})}{\sigma\sqrt{m}} \ , \quad 0\le t\le 1,$$

where $\lfloor\cdot\rfloor$ is the greatest integer function. $\{T_m(t),\ 0\le t\le 1\}$ is called the *standardized time series*. Schruben (1983) proves that if $X_1,\ldots,X_m$ is a stationary, φ-mixing, finite variance sequence of random variables (satisfying one other technical condition), then as $m\to\infty$, $T_m(t)$ converges in distribution to a *standard Brownian bridge* process, $\{B_t,\ 0\le t\le 1\}$. Also, the standardized time series is asymptotically independent of $m\bar{X}_m$.

<u>Remark</u>: It is well known that $B_t \sim \mathrm{Nor}(0,t(1-t))$ and $\mathrm{Cov}(B_{t_1},B_{t_2}) = \min(t_1,t_2)(1-\max(t_1,t_2))$.

Define $A \equiv \frac{\sigma}{m}\sum_{k=1}^{m} c_k T_m(k/m)$, the c_k's being pre-specified. For large m, $A \approx$ (is approximately distributed as) $\sigma\mathrm{Nor}(0,V)$, where

$$V \equiv \frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m} \mathrm{Cov}(c_j B_{j/m}, c_k B_{k/m}) \ .$$

It will be computationally convenient to approximate V by letting $j = ms$ and $k = mt$ (so that $djdk = m^2 dsdt$). Thus,

$$V = \frac{1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} c_j c_k \min(j/m,k/m)[1-\max(j/m,k/m)]$$

$$\doteq \int_0^1 \int_0^1 c_{ms} c_{mt} \min(s,t)[1-\max(s,t)]\, dsdt$$

$$= 2 \int_0^1 \int_0^t c_{ms} c_{mt}\, s(1-t)\, dsdt \,. \tag{1}$$

So $A^2/V \doteq \sigma^2 \chi^2(1)$, and this is called the *weighted area estimator* for the variance σ^2.

Suppose now that we work with the process $X_1,\ldots,X_n$, where $n=bm$, and that this series satisfies Schruben's conditions. Divide the process into b contiguous *batches*, each of size m; i.e., $X_{(i-1)m+1}$, $X_{(i-1)m+2}$, ..., X_{im} comprise batch i, $i=1,\ldots,b$. Each individual batch can be standardized: For $i=1,\ldots,b$ and $j=1,\ldots,m$, let

$$\bar{X}_{i,j} \equiv \frac{1}{j} \sum_{k=1}^{j} X_{(i-1)m+k} \qquad \text{(average of the first } j\ X\text{'s from the } i\text{-th batch)},$$

$$\bar{\bar{X}}_n \equiv \frac{1}{n} \sum_{k=1}^{n} X_k \quad \text{(grand mean)},$$

$$T_{i,m}(t) \equiv \frac{\lfloor mt \rfloor (\bar{X}_{i,m} - \bar{X}_{i,\lfloor mt \rfloor})}{\sigma\sqrt{m}}\,, \quad 0 \le t \le 1, \quad \text{and}$$

$$A_i \equiv \frac{\sigma}{m} \sum_{k=1}^{m} c_k T_{i,m}(k/m) \,.$$

For large enough m, each of the standardized time series [the $T_{i,m}(t)$'s] is approximately distributed as a Brownian bridge; so $A_i \doteq \sigma \mathrm{Nor}(0,V)$, $i=1,\ldots,b$. Further, for large m, we can treat the batches as if they were (approximately) independent. This yields:

$$\frac{1}{V}\sum_{i=1}^{b} A_i^2 \;\Rightarrow\; \sigma^2\chi^2(b) .$$

An immediate consequence of Theorem 21.1 of Billingsley (1968) is that

$$Z_n \equiv \frac{\bar{\bar{X}}_n-\mu}{\sigma/\sqrt{n}} \;\Rightarrow\; \mathrm{Nor}(0,1) .$$

Then the asymptotic independence result above gives:

$$\frac{Z_n}{\left[\dfrac{\Sigma A_i^2}{\sigma^2 bV}\right]^{1/2}} \;\Rightarrow\; \frac{\mathrm{Nor}(0,1)}{\left[\dfrac{\chi^2(b)}{b}\right]^{1/2}} \sim t(b) .$$

We finally obtain confidence interval estimators for μ:

$$\Pr\left\{ \mu \in \bar{\bar{X}}_n \pm t_{b,1-\alpha/2}\left[\frac{1}{nbV} \sum_{i=1}^{b} A_i^2 \right]^{1/2} \right\} \;\Rightarrow\; 1-\alpha , \qquad (2)$$

where $t_{b,\nu}$ is the upper-ν quantile of the $t(b)$ distribution.

We consider various choices for the weights (the c_k's). These choices and their resulting V's [from the integral approximation (1)] are summarized in Table 1.

Remarks:

(i) After choosing the weighting sequence $\{c_j\}$, the associated V from Table 1 is used in (2) to form a confidence interval estimator.

(ii) The calculation of V from (1) is straightforward (but sometimes tedious).

Example: For choice 3 (from Table 1),

$$V = 2\int_0^1\int_0^1 \frac{1}{s(1-s)}\,\frac{1}{t(1-t)}\, s(1-t)\, ds\, dt$$

$$= 2\int_0^1\int_0^t \frac{1}{(1-s)t}\, ds\, dt \;=\; -2\int_0^1 \frac{\ln(1-t)}{t}\, dt \;=\; \frac{\pi^2}{3}\,.$$

(iii) The variance estimator which results from choice 1 (the equal weighting case) is asymptotically the same (as $m\to\infty$) as the so-called *area* variance estimator from Schruben (1983).

(iv) For each standardized time series, $\{T_{i,m}(m)\}$, choice 2 grants greater weight for 'small' values of t. Choices 3 through 6 give comparatively little weight to the middle ($t \approx 1/2$) of each standardized time series.

(v) Table 2 summarizes an empirical study involving the order 1 exponential autoregressive model [cf. Lewis, (1980)]. The weighted area estimators are seen to perform well for 'large' batch size.

(vi) Denote the random variable corresponding to the half-length of the weighted area estimator by H. Following Schmeiser (1982) and Goldsman and Schruben (1984) (G-S), it is easy to derive the following:

$$E[H] = \frac{\sigma}{\sqrt{n}}\, t_{b,1-\alpha/2}(2/b)^{1/2}\,\frac{\Gamma((b+1)/2)}{\Gamma(b/2)}\,,$$

$$Var(H) = \frac{\sigma^2}{n}\, t^2_{b,1-\alpha/2}\left\{1 - \frac{2}{b}\left[\frac{\Gamma((b+1)/2)}{\Gamma(b/2)}\right]^2\right\}, \text{ and}$$

The coverage probability,

$$Pr\{|\bar{\bar{X}}_n-\mu_1| < H\} = F(t_{b,1-\alpha/2}) - F(-t_{b,1-\alpha/2})\ , \text{ where}$$

$\Gamma(.)$ is the gamma function and F(.) is the c.d.f. of the noncentral

of correlation amongst batches which are encountered when using the area estimator.

References

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons, New York.

Goldsman, D. (1984). *On Using Standardized Time Series to Analyze Stochastic Processes.* Ph.D. Thesis, School of OR&IE, Cornell University, Ithaca, NY.

Goldsman, D. and L. Schruben (1984). Asymptotic properties of some confidence interval estimators for simulation output. *Management Science*, 30, 1217-1225.

Lewis, P.A.W. (1980). Simple models for positive-valued and discrete-valued time series with ARMA correlation structures. From *Multivariate Analysis-V*, pp. 151-166, P.R. Krishnaiah (ed.) North Holland, Amsterdam.

Schmeiser, B. (1982). Batch size effects in the analysis of simulation output. *Operations Research*, 30, 556-568.

Schruben, L. (1983). Confidence interval estimation using standardized time series. *Operations Research*, 31, 1090-1108.

Table 1 - Choices of Weights and Resulting V's

weighting choice	c_j, $j=1,\ldots,m$	V from (1)
1	1	$1/12$
2	$1-\frac{j}{m}$	$1/45$
3	$\left[\frac{j}{m}\left(1-\frac{j}{m}\right)\right]^{-1}$	$\frac{\pi^2}{3}$
4	$\left\lvert\frac{1}{2}-\frac{j}{m}\right\rvert+\epsilon \quad (\epsilon\geq 0)$	$\frac{1}{320}+\frac{\epsilon}{32}+\frac{\epsilon}{12}$
5	$\left[\frac{1}{2}-\frac{j}{m}\right]^2+\epsilon \quad (\epsilon\geq 0)$	$\frac{1}{4032}+\frac{\epsilon}{120}+\frac{\epsilon^2}{12}$

Table 2

Performance of weighted area confidence interval estimators for the mean of an EAR(1) process with coefficient ρ = 0.2 and exponential (mean=1) noise based on 100 independent runs of 2560 observations each. [Choices of weights are summarized in Table 1].

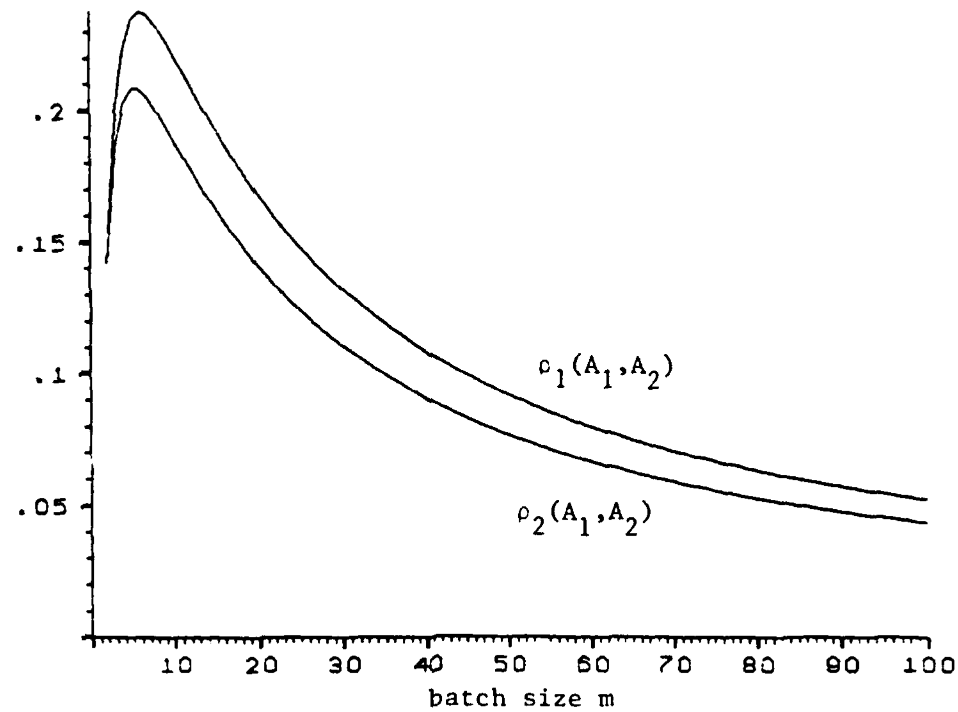
b	m \ weighting choice	1	2	3	4(ϵ=0)	5(ϵ=0)
	Confidence interval achieved coverage (90% desired)					
1	2560	.88	.92	.90	.89	.90
2	1280	.93	.96	.94	.93	.95
5	512	.95	.94	.96	.95	.96
10	256	.93	.93	.94	.95	.96
20	128	.93	.94	.91	.91	.91
	Average confidence interval half-length (x 10000)					
1	2560	1193	1266	1184	1199	1202
2	1280	592	612	598	613	617
5	512	480	481	486	492	498
10	256	440	441	436	444	446
20	128	426	423	415	427	426
	Sample standard deviation of half-lengths (x 10000)					
1	2560	900	923	876	932	931
2	1280	297	300	298	294	298
5	512	153	149	150	150	149
10	256	103	104	97.5	95.6	93.7
20	128	52.6	50.7	57.9	59.5	65.5

Table 3 - Typical Small-Sample Values of ρ_1 and ρ_2

(m = batch size)

α	m	ρ_1	ρ_2
0.5	15	-0.04268	-0.03686
0.5	20	-0.03231	-0.02767
0.5	25	-0.02600	-0.02215
0.0	--	0	0
-0.5	15	0.19445	0.16427
-0.5	20	0.16964	0.14261
-0.5	25	0.14969	0.12550
-0.9	15	0.42568	0.35214
-0.9	25	0.44111	0.36053
-0.9	50	0.43968	0.35646
-0.9	100	0.41352	0.33456

Figure 1: Correlation of A_1, A_2 versus batch size m for an MA(1) process with $\alpha = -0.5$



REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS
Unrestricted

2a. SECURITY CLASSIFICATION AUTHORITY
Office of Naval Research

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE
Not applicable

3. DISTRIBUTION/AVAILABILITY OF REPORT
Unlimited Distribution

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Technical Report No. J-84-15

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Cornell University

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
Office of Naval Research

6c. ADDRESS *(City, State and ZIP Code)*
Ithaca, NY 14853

7b. ADDRESS *(City, State and ZIP Code)*
Arlington, VA 22217

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
Office of Naval Research

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-81-K-0037

8c. ADDRESS *(City, State and ZIP Code)*
Arlington, VA 22217

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE *(Include Security Classification)*
New confidence interval estimators using standardized time series (unclassified)

12. PERSONAL AUTHOR(S)
David Goldsman and Lee Schruben

13a. TYPE OF REPORT
Interim

13b. TIME COVERED
FROM Mar 31 84 TO Aug 25 85

14. DATE OF REPORT *(Yr., Mo., Day)*
1984 December

15. PAGE COUNT
10

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

We develop new confidence interval estimators for the underlying mean of a stationary simulation process. These estimators can be viewed as generalizations of Schruben's so-called *standardized time series area* confidence interval estimators. Various properties of the new estimators are given.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Lee W. Schruben

22b. TELEPHONE NUMBER *(Include Area Code)*
(607) 256-4856

22c. OFFICE SYMBOL

END

FILMED

1-86

DTIC